11-30-2020

Publishing E-resources of Digital Institutional Repository as Linked Open Data: an experimental study

Subhendu Kar
Vivekananda Satavarshiki Mahavidyalaya, subhendu.kar@gmail.com

Dr. Rajesh Das
The University of Burdwan, rajeshdas99@gmail.com

Publishing E-resources of Digital Institutional Repository as Linked Open Data: an experimental study

Subhendu Kar, Librarian, Vivekananda Satavarshiki Mahavidyalaya, Manikpara, Dist, Jhargram, West Bengal, India Pin-721513
And
Dr. Rajesh Das, Assistant Professor, Department of Library and Information Science, The University of Burdwan, Burdwan, West Bengal, India, Pin-713104

Abstract

Linked open data (LOD) is an essential component in semantic web architecture and is becoming increasingly important over time due to its ability to share and re-use structured data which is both human and computer readable over the web. Currently, many libraries, archives, museums etc. are using open source digital library software to manage and preserve their digital collections. They may also intend to publish their e-resources as "Linked Open Datasets" for further usage. LOD enables the libraries or information centers to publish and share the structured metadata that is generated and maintained with their own bibliographic and authority data in such a way that the other libraries and general community across the world can consume, interact, enrich and share. In this context, the key issue is to convert the library bibliographic data which is commonly known as metadata into LOD dataset. The purpose of this paper is to provide a methodology and technical aspects to design and publish a structured LOD dataset of bibliographic information from a digital repository developed with DSpace digital library software so that other libraries can link their repositories with these LOD for providing additional relevant resources to their end-users. The paper shows the process of integration and configuration of Apache Jena Fuseki (a tool for SPARQL Endpoint interface) with DSpace for converting metadata into Resource Description Framework (RDF) triple model and make them available in various RDF formats. It also discusses a model for building a LOD framework to convert and store RDF graph and RDF triple. Finally, it tests the accessibility of the inked open dataset by querying RDF data through a SPARQL endpoint interface.

Keywords: Linked Data (LD), Linked Open Data (LOD), Resource Description Framework (RDF), SPARQL Endpoint, DSpace, Apache Jena Fuseki, Ontology, Metadata Framework, Metadata Mapping

1 Introduction

Over a decade, Linked Open Data (LOD) has gained immense popularity as a way of sharing and re-using structured data over the World Wide Web since the inception of the concept of Linked Data (LD) in 2006. Heath and Bizer (2011) in their book proposed LOD as a framework for the semantic web, as it enables us to handle data from various disciplines with a variety of granularity. Kaltenböck and Bauer (2012), referred LOD as a growing trend for organization to provide a machine-readable format for their existing data which make easier for users to use the existing datasets, create and merge the datasets and make available their own datasets in shareable formats and applications. They have also mentioned some essential criteria and principles for the data to become 'open data' in the Linked Data context. LOD provides substantial improvements over conventional methods of managing, processing and disseminating library data by offering a streamlined functional alternative of the collaborative sharing mechanisms in digital library environment. Bibliographic data traditionally managed by libraries and other information centers are commonly defined as metadata in digital archiving or preservation. In the context of LOD, managing and publishing the metadata in a structured and systematic manner is the biggest concern. During the creation and publication of library data into LOD, the libraries need to follow the design principles of linked data and five-star schema of LOD as proposed by Tim Berners-Lee in 2006. During the last few years many libraries, archives, museums across the globe have been trying to publish LOD datasets of the resources of their digital repositories. In this paper we have attempted to design a structured LOD dataset of bibliographic information from the digital repository containing more than 500 bibliographic data. The repository which has been developed with the open source software DSpace, contains almost five hundred full-text bibliographic items most of which are in-house and open access publications. A series of process converting the data into RDF triple and publishing them in a user-friendly interface are the key aspects of this work.

2 Review of Literature

In the context of semantic web or web of data, Bizer, Heath and Berners-Lee (2009) explain the basic concept and design principles of linked data. The authors emphasize on the recent development of related technologies in the area of semantic web. They also described the development in publishing linked data on the web and also reviewed some major applications developed to utilized the web of data. They outlined a research agenda for the Linked Data community to move the development forward.

Heath and Bizer (2011) in their book proposes LOD as a framework for the semantic web, as it enables us to handle data from various disciplines with a variety of granularity. Mitchell (2013) provides a thorough concept of linked open data and linked open vocabularies on the basis of five building blocks of linked open metadata viz data model, data serialization, metadata schema, content rules and data exchange. In their article, Roy Chowdhury and Das (2019) provides a theoretical approach to describe the benefits of implementing LOD in libraries with reference to some major initiatives worldwide and highlights the possible challenges in adopting LOD in libraries. Zengenene (2013) provides an overview of the concept of Linked Data and explained how libraries would be benefited from it. The paper found that Linked Data has become a topic of interdisciplinary interest. A study undertaken by Hallo and others (2015) explores the current status and best practices in the implementation of LOD in digital libraries across the globe through a literature search and information from the websites of the projects under study. The research highlights the use of vocabularies and ontologies; and attempts to find the advantages and issues in implementing linked data in digital libraries. An experiment by Papadakis and Kyprianos (2012) explores the relevance of subject-based browsing services based on LOD created with the bibliographic data of a digital repository built with DSpace and examines its capability of linking resources from multiple repositories. They observe that LOD services for open data are still lacking towards the end-user. LOD allows libraries to reuse data created by other libraries or information centers with reference to some cases of conversion of bibliographic and authority data into LOD by the Library of Congress, the German National Library, and the Swedish National Library (Borst, Fingerle, Neubert, & Seiler, 2010). Halla (2013) discusses the application of Linked Data principles on cataloging based on Library of Congress Bibliographic Framework, BIBFRAME, a data model for bibliographic description through an extensive literature search published in the Library and Information Science domain. The purpose of the research paper is to evaluate the applicability of BIBFRAME in publishing cataloging data as linked data and examines whether it is useful to overcome the drawbacks of MARC-based cataloging as discussed in the literature under review. The paper provides an exhaustive concept of linked data and linked open data as published in literature and the significance of publication and consumption of LOD in library catalog. This case study discusses the conversion and publication of serial publication into LOD using WordPress Content Management System and mapping with Dublin Core Metadata standards. This study is based on the full-text articles and proceedings of the official medical research journal of Shaikh Zayed Medical Complex, Pakistan (Basit & Hussain, 2019). A research from Konstantinou, Houssos, and Manta (2014) shows the process of generating bibliographic information of digital repositories as linked open data using ontology mapping based on international standards. Their study focuses on integration, expressiveness and query answering capability in the implementation of linked open data in a methodical way. Khalili, Loizou, and Harmelen (2016) find the significance and demand of user-oriented linked data applications that require knowledge of semantic web technologies like RDF, ontology vocabulary, metadata schema, SPARQL query language to design a user-centrist interface for viewing, editing and browsing linked data. Egusa and Takaku (2020) conduct a case study for designing and publishing a LOD dataset of bibliographic data for Japanese Textbook at the National Institute for Educational Policy Research. The dataset consists of RDF triples of bibliographic information of more than seven thousand Japanese textbooks This paper shows the process of publishing the LOD dataset online by assigning Uniform Resource Identifier (URI) to each item and designing a metadata model for the same.

3 Statement of Problem

The research work has been carried out to explore how the bibliographic data of a library can be transformed and published as linked open data (LOD) by following a specific RDF (Resource Description Framework) metadata schema through the enrichment of external data resources in a digital library environment. This is the main problem of the research work that can be further divided into some specific questions. The research questions are discussed in the following section.

3.1 Research Questions

The research questions that have emerged in relation to this research work are:

- i Is it possible to publish the bibliographic resources of a library as linked open data in a digital library environment?
- ii Is it possible to provide linked open data service in a digital library platform that uses a free and open source (FOSS) digital library or institutional repository (IR) software?
- iii Is it possible to transform all sorts of bibliographic resources of a digital library into linked open dataset in RDF model?
- iv Can linked open datasets be provided as RDF triple model with the integration of domain specific RDF metadata schema in a digital library environment?
- v Can it be possible to design and implement a SPARQL (SPARQL Protocol and RDF Query Language) endpoint interface for querying linked open dataset?

4 Objectives

The primary objective of this research work is to design and implement an interface for publishing the linked data services for openly accessible library resources in a digital library environment. To be more specific, the goals of this research work can be pointed out as follows:

i To publish the library bibliographic resources as linked open data in a digital library environment.

ii To publish linked open data through a FOSS based institutional repository (IR) software in a digital library environment.

iii To convert all existing library bibliographic resources available in the digital repository under study as RDF Triple or linked open dataset.

iv To provide the digital library resources as linked open data with integration of domain specific RDF metadata schema.

v To integrate SPARQL Endpoint interface for querying and accessing RDF triple dataset in a digital library environment.

5 Importance of linked open data in libraries

Bibliographic information that is converted into structured metadata and published as LOD by the library and information centers across the world, can be shared in such a manner that the other libraries or general community can consume, interact, share and enrich the data. LOD allows assigning suitable unique identifiers corresponding to the types of data that a library, or an information center, or an archive, or a museum contains; publishes them using standard ontology languages for mapping the metadata for efficient resource discovery and sharing over the web. It facilitates the publication of structured data that is both machine-readable and human-readable with an open license and makes them accessible and downloadable through various open formats and it allows the exchange of the structured data to re-use by the other libraries or institutions. LOD enhances the visibility of the library data and helps in establishing links with other online services. It enables multiple queries on linked metadata from multiple points and also improves the credibility of structured data by allowing end users' annotations to the resources.

6 Methodology

In this research, a set of methods have been followed to perform the research work. The research methodology has been divided into two parts:

i Activities involving the aspects of a free and open-source digital library software; and i Activities include transforming all the existing resources of a digital library into linked open dataset following a domain-specific RDF metadata schema and designing supporting SPARQL queries.

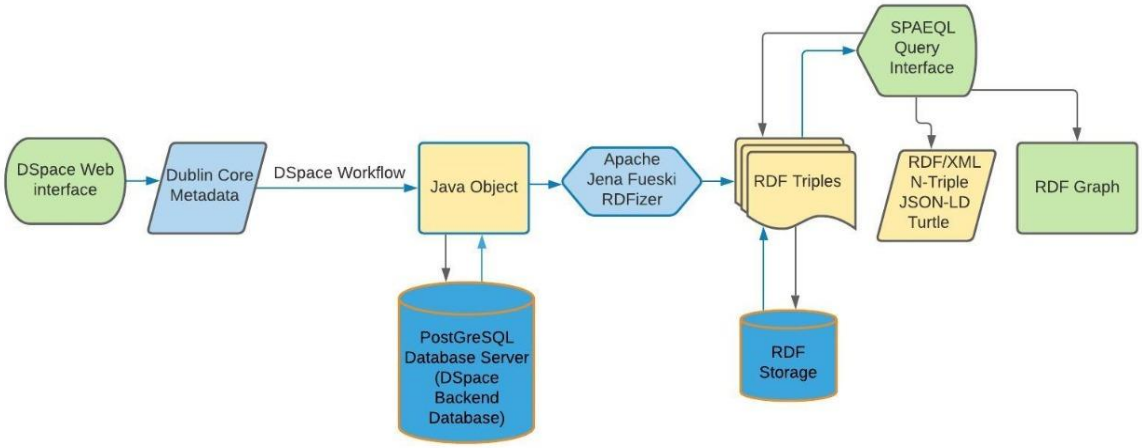


Figure-1: Conceptual Framework for transforming of full-text resources metadata to RDF triples

Here, a conceptual framework has been drawn for interpreting the activities involved in the workflow of conversion and publication of bibliographic data into RDF triples at a glance. In Figure-1, it is clearly shown how metadata could be transformed to RDF triple data by using open source software and tools. In this experimental study, DSpace (https://duraspace.org/dspace/) software has been used as a digital library system and Apache Jena Fuseki (https://jena.apache.org/) has been used for RDF data manipulation and queries. DSpace uses PostgreSQL(https://aws.amazon.com/rds/postgresql) as a back-end database server and follows Dublin Core Metadata Schema for resource description. In the DSpace system, metadata is stored in the PostgreSQL database. To convert these data into RDF triples, we used RDFizer as a converter included in the Apache Jena Fuseki software. From the above diagram, we can see how the data stored in the PostgreSQL database is converted in RDF triples, stored into RDF storage area, retrieved as RDF format, and visualized as RDF graph.

6.1 Specific Steps in the Methodology

6.1.1 Experimental data collection

The collection of required data has been made for this research work from a digital repository of a general degree college, V.S. Mahavidyalaya in West Bengal, India. This institutional repository has been developed with a free and open source digital library software, DSpace. The repository contains almost five hundred full-text bibliographic items most of which are in-house publications. The resources of this repository are intended to be used for academic and research purposes.

6.1.2 Identification and Selection of Digital Library Software

At present, free and open source digital library software packages are being widely implemented in different sectors for digital archiving. In terms of design, architecture, and framework, many open source digital library applications are very efficient and secured system like any proprietary software. Nowadays, we can find a range of free and open source applications in a digital archive or digital library domain, such as DSpace, Greenstone Digital Library (GSDL) (http://www.greenstone.org/), Eprints (https://www.eprints.org/uk/), Fedora Commons (https://duraspace.org/fedora/), Omeka (https://omeka.org/), CONTENTdm (https://www.oclc.org/en/contentdm.html), and many more. We have made a comprehensive study in OpenDOAR directory (https://v2.sherpa.ac.uk/opendoar) to find out the applicability and usability of different types of digital library software. As a result, we have found, over sixty percent of repositories are using DSpace digital library software. GSDL, Fedora Commons, and Omeka lead the table after DSpace. DSpace is based on Java architecture and it is a very secured software with a good level of authentication and encryption. Not only that, it uses Apache Solr Search server that runs on the top layer of Lucene search engine. DSpace has a powerful indexing system and it is suitable for hierarchical collection development for an organization. DSpace releases frequent updates to its versions, thereby maintaining the latest technological changes. Among the digital library software mentioned above, DSpace can only support the LOD service. For these reasons, we have selected DSpace software for the creation of the digital repository.

6.1.3 Selection of LOD Architecture

This research work aims to publish the textual contents stored in a digital library as linked open data that has to be converted into an RDF Triple model. After the implementation of DSpace, it needs to build a LOD architecture or framework to convert and store the contents into a structured dataset (RDF graph and RDF triple). After the development and integration of DSpace with LOD, an RDF triple store has to be installed and configured for accessing converted DSpace data as RDF triple format like RDF/XML, Turtle, N-Triple, JSON-LD, etc. There are several SPARQL (RDF) query interfaces available, such as OpenVirtusa, Apache Jena Fuseki, etc. DSpace software can work with any of these above mentioned SPARQL endpoint interface.

For this work, Apache Jena Fuseki version 3.16 has been implemented for LOD query interface. We can directly invoke the DSpace repository from this SPARQL Endpoint interface (Please see DSpace official documentation for details). DSpace allows configuring the RDF module for accessing the LOD dataset. DSpace, version 5.0 onward supports integration of LOD and RDF Triple. RDF Triple data can be serialized into various RDF formats like RDF/XML, Turtle, N-Triple, JSON-LD, and many others. Six main files are located in the DSpace source directory, (detail of which will be discussed later), viz. *dspace.cfg', 'rdf.cfg', 'constant-data-*.ttl', 'metadata-rdf-mapping.ttl', 'fuseki-assembler.ttl'* and *'rdf.xml'*. Apart from the files mentioned above, other associated files will be configured as per requirements. The main DSpace configuration file *(dspace.cfg)* and the RDF file *(rdf.cfg)* are primarily configured for LOD support. For this research work, the primary configuration of ds*pace.cfg* and *rdf.cfg,* is performed following the main DSpace configuration (https://www.lyrasis.org/DCSP/Pages/DSpace.aspx).

6.1.4 Installation and Configuration of SPARQL Endpoint Interface

As said earlier, there are many SPARQL endpoint interfaces available for query language and RDF graph store to access and visualize the RDF triple data. These two applications work with DSpace to support DSpace RDF service. We have installed and configured Apache Jena Fuseki Server version 3.16 for converting the bibliographic data of the DSpace repository as RDF triple data. Apache Jena Fuseki is a semantic web framework for RDF data and RDF graphs that supports SPARQL query language and SPARQL graph store. Fuseki is an HTTP interface for accessing RDF data. (https://jena.apache.org/documentation/fuseki2/).

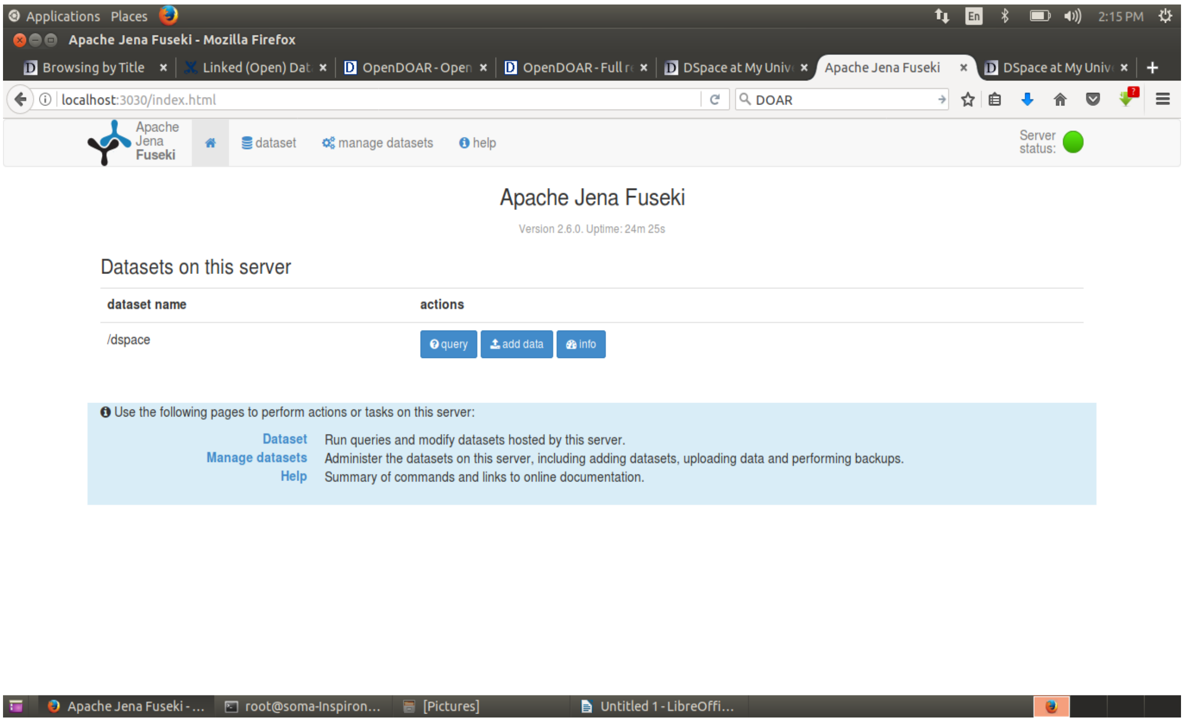


Figure-2 : Default Homepage of Apache Jena Fuseki v. 3.16

6.1.5 Mapping of DSpace metadata and RDF vocabulary

A web resource or web content comes with metadata. The metadata allows to describe content of the resources. Simply metadata refers data about data. There are many metadata schema available for describing the resources which are available on the web. Dublin Core (DC) metadata schema is one of the most popular general domain metadata schema for describing e-resources on the Internet. DSpace repository follows Dublin Core metadata schema as its default metadata schema. If we want to build a repository with the content of general bibliographic e-resources, we follow the Dublin Core metadata schema as default. If we want to develop domain-specific digital repositories like learning resources repository on a specific discipline, a repository of cultural heritage resources, repository comprising resources of archive and museums, the institutional repository of e-thesis and dissertation, etc., we need to customize the Dublin Core metadata schema and its elements as per our requirements in the DSpace system. In DSpace digital repository system, customization of Dublin Core metadata schema and elements is very flexible and logical to build a domain specific digital library system or digital archive. We know that every e-resource comes with a metadata schema in any of a variety of formats, such as database, table, excel, XML, CSV, and so on. So before conversion of DSpace items into RDF triple format, we have to assign an RDF metadata schema and RDF vocabulary with converted DSpace data. The process of metadata mapping defines how to map a specific metadata filed within DSpace to a triple that will be added to the converted data. The main configuration file *'rdf.cfg'* contains the property *'MetadataConverterPlugin'* that uses the DSpace metadata mapping to convert the metadata of an item into RDF. The main file of mapping between DSpace metadata and RDF is *'metadata-mapping.ttl'*. In this research work, the DC metadata of DSpace has been mapped with 'DSpace Metadata RDF Mapping Vocabulary' (http://digital-repositories.org/ontologies/DSpace-metadata-mapping/). This mapping vocabulary is a simple ontology that includes properties, classes, sub-classes, and relationships. Before converting the DSpace data, every DC metadata value is to be assigned with either one of the following triple viz., subject, predicate, and object. Subject and predicate should be replaced by *'DSpaceObjectIRI'* of the mapping vocabulary. The object part of the triple is to be assigned with a value as 'literal'. The namespace of the DSpace Metadata Mapping Vocabulary is presented as 'dm' which is used in the main configuration file. The following table represents the mapping between Dublin Core metadata elements and DSpace metadata RDF mapping vocabulary.

Metadata	**Dublin Core metadata elements**	**RDF fields (Subject/Predicate/Object)**
Title	dc.title	dm:subject --> dm:DSpaceObjectIRI dm:predicate--> dcterms:title dm:object--> dm:DSpaceValue
Title: Alternative	dc.title.alternative	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> dcterms:alternative dm:object--> dm:DSpaceValue
Contributer: Author	dc.contributor.author	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> dc:creator dm:object--> dm:DSpaceValue
Contributor	dc.contributor	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> dc:contributor dm:object--> dm:DSpaceValue
Date: Available	dc.date.available	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> dcterms:available dm:object--> a dm:LiteralGenerator dm:pattern "$DSpaceValue" dm:literalType xsd:dateTime
Date: Copyright	dc.date.copyright	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> dcterms:dateCopyrighted dm:object--> a dm:LiteralGenerator dm:pattern "$DSpaceValue"

		dm:literalType xsd:dateTime
Date: Created	dc.date.created	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> dcterms:created dm:object--> a dm:LiteralGenerator dm:pattern "$DSpaceValue" dm:literalType xsd:dateTime
Date: Issued	dc.date.issued	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> dcterms:issued dm:object--> dm:DSpaceValue
Date: Submitted	dc.date.submitted	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> dcterms:dateSubmitted dm:object--> a dm:LiteralGenerator dm:pattern "$DSpaceValue" dm:literalType xsd:dateTime
Date: Updated	dc.date.updated	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> dcterms:modified dm:object--> a dm:LiteralGenerator dm:pattern "$DSpaceValue" dm:literalType xsd:dateTime
Date	dc.date	dm:subject-->dm:DSpaceObjectIRI dm:predicate--> dc:date dm:object--> a dm:LiteralGenerator dm:pattern "$DSpaceValue" dm:literalType xsd:dateTime
Description:Abstract	dc.description.abstract	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> dcterms:abstract dm:object--> a dm:LiteralGenerator dm:pattern "$DSpaceValue" dm:DSpaceLanguageTag "true"^^xsd:boolean
Format	dc.format	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> dc:format dm:object--> dm:DSpaceValue
Identifier: URI	dc.identifier.uri	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> bibo:doi dm:object--> a dm:LiteralGenerator dm:modifier [dm:matcher "^http://dx.doi.org/(.*)$"

		dm:replacement "doi:$1"]
Identifier: Citation	dc.identifier.citation	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> dcterms:bibliographicCitation dm:object--> dm:DSpaceValue
Identifier: ISBN	dc.identifier.isbn	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> bibo:issn dm:object--> dm:DSpaceValue
Itenfier:ISSN	dc.identifier.issn	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> bibo:issn dm:object--> dm:DSpaceValue
Language	dc.language	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> dc:language dm:object--> dm:DSpaceValue
Publisher	dc.publisher	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> dc:publisher dm:object--> dm:DSpaceValue
Rights	dc.rights	dm:subject--> dm:DSpaceObjectIRI dm:predicate--> dc:rights dm:object--> dm:DSpaceValue

Table-1: Mapping between Dublin Core Metadata Elements and DSpace Metadata RDF Mapping Vocabulary

DSpace uses Dublin Core (DC) metadata schema as the default metadata schema. The description of items or objects in DSpace is supported by the DC metadata schema. DC metadata schema describes the web resources in general. This metadata schema can be used by any web resource in the general domain for tagging and describing the data elements for better resource discovery. But in the DSpace system, the administrator may wish to customize the metadata schema as per his own need. For example, it could be customized for processing electronic theses and dissertations (ETDs), cultural heritage resources, learning object resources, moving image resources, and many more. On the other hand, RDF triple data could be followed by RDF metadata schema/ontology/vocabulary. RDF metadata schema is a simple metadata schema which describes the RDF data field like subject, predicate and object. The ontologies and vocabularies not only describe the RDF data, but also used for presenting the relationship between the RDF data. Table 1 represents the mapping between the elements of DSpace metadata schema and RDF field. Apart from these, the relationship between RDF fields is also present in the mapping configuration file (*'metadata-mapping.tpl'*) in the DSpace repository. We

can use more than one schema or ontology or vocabulary for highly structured RDF data representation. In conversion stage, DSpace-RDF configuration includes a list of prefix that describe and establish the relationship between RDF data.
A list of prefix with their URI (namespace) have been shown in following table:

Prefix	**Name**	**URI (Namespace)**
rdf	The RDF Concepts Vocabulary	http://www.w3.org/1999/02/22-rdf-syntax-ns#
rdfs	The RDF Schema vocabulary	http://www.w3.org/2000/01/rdf-schema#
owl	Web Ontology	http://www.w3.org/2002/07/owl#
xsd	XML Schema	http://www.w3.org/2001/XMLSchema#
dc	DCMI Metadata Terms, v 1.1	http://purl.org/dc/elements/1.1/
dcterms	DCMI Metadata Terms	http://purl.org/dc/terms/
bibo	Bibliographical Ontology	http://purl.org/ontology/bibo/
dm	DSpace Metadata RDF mapping vocabulary	http://digital-repositories.org/ontologies/DSpace-metadata-mapping/0.2.0#

Table-2: Namespaces and Vocabularies used in RDF dataset

6.1.6 Conversion of DSpace Metadata to RDF Triple Data

The DSpace uses PostgreSQL database server as back-end database server of the repository. This step includes the conversion of DSpace data into RDF triple as well as RDF graph. That will be accessed through Apache Jena Fuseki SPARQL interface. DSpace allows a single line command for this conversion of entire DSpace items into RDF data. All data is to be converted as RDF triple format with the provenance of subject, predicate, and object. For this research work, we have used the command *'/DSpace/bin/DSpace rdfizer --convert-all once'* to convert the DSpace data into RDF. The repository under study contains almost five hundred full-text bibliographic items of various types. The metadata of all items are converted into RDF triple.

7 Evaluation

The main purpose is to provide the LOD services in a digital library environment with the conversion of all full-text resources into linked open data. For this research work, more than 500 full-text items have been converted into RDF triples from a DSpace repository. In the methodology section, we have demonstrated how to install a third party SPARQL endpoint and how DSpace metadata can be converted into RDF triple by using RDFizer. Apache Jena Fuseki SPARQL endpoint interface has been used for querying RDF data. After completion of the conversion process of DSpace metadata into RDF triples, a URI of DSpace RDF SPARQL (http://localhost:3030/DSpace/sparql) is found for making SPARQL queries of DSpace data. A simple SPARQL statement has been tested in the SPARQL interface by using various prefixes that are shown in figure 3. Here, we can see the statement for querying all the data from DSpace-RDF triples as subject, predicate and object {?s ?o ?p}. We can see the output as RDF triple format (subject, predicate, object) in figure 4 with URIs and literals. Subject must always contain the URI but object and predicate may contain URI and/or literal both.

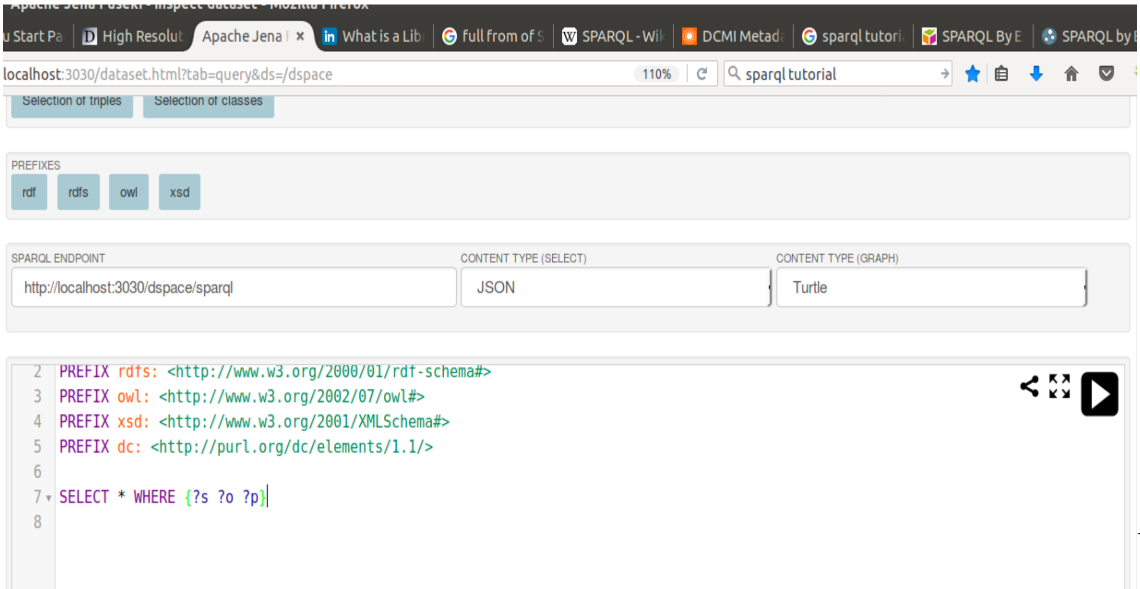


Figure-3: SPARQL query statement with prefixes

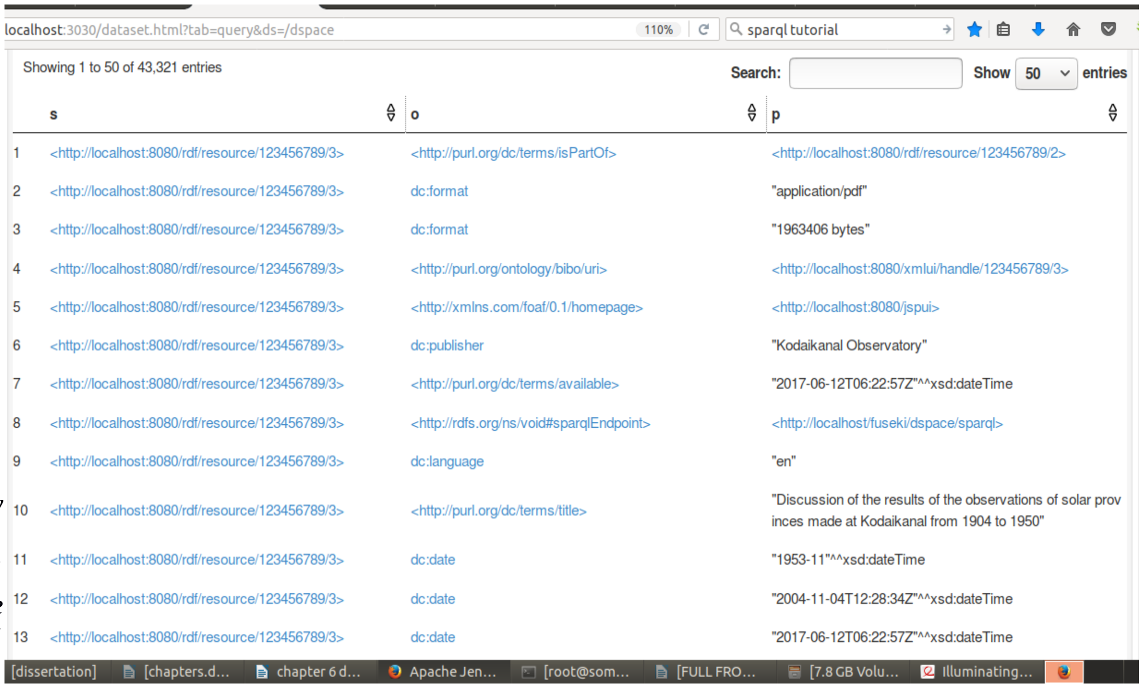
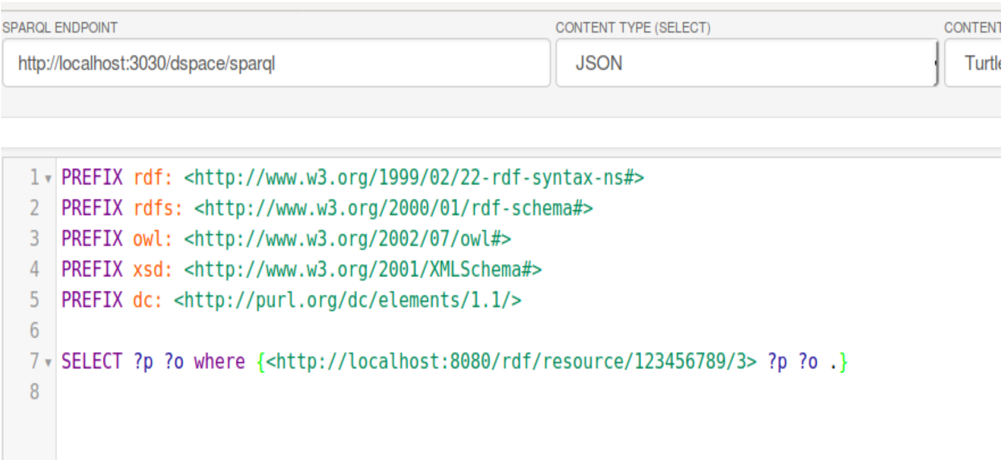


Figure-4: Showing RDF triple results

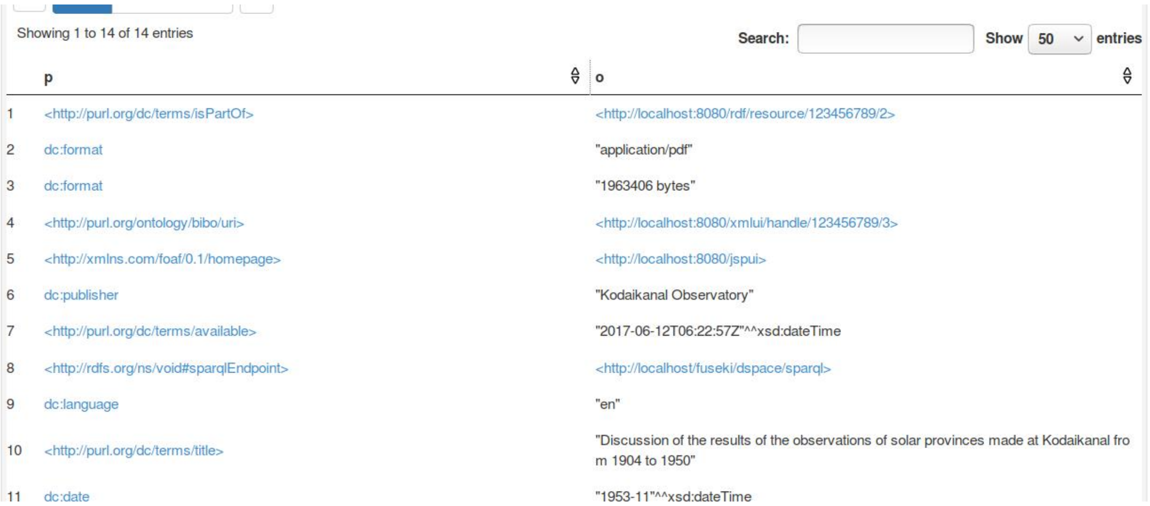



```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX dc: <http://purl.org/dc/elements/1.1/>

SELECT ?p ?o where {<http://localhost:8080/rdf/resource/123456789/3> ?p ?o .}
```

Figure-5: SPARQL query statement for a particular item

In figure 5, another statement that denotes a particular URI of a subject element is written as ‘http://localhost:8080/rdf/resources/123456789/3’. As a result, we can see the predicate and object of the RDF item (in figure 6). We can also use the SPARQL statement to make queries about the RDF graph. We may store the output data of a specific item or group of items in different formats, such as RDF / XML, N-triple, JSON-LD, Turtle, etc.



	p	o
1	<http://purl.org/dc/terms/isPartOf>	<http://localhost:8080/rdf/resource/123456789/2>
2	dc:format	"application/pdf"
3	dc:format	"1963406 bytes"
4	<http://purl.org/ontology/bibo/uri>	<http://localhost:8080/xmlui/handle/123456789/3>
5	<http://xmlns.com/foaf/0.1/homepage>	<http://localhost:8080/jspui>
6	dc:publisher	"Kodaikanal Observatory"
7	<http://purl.org/dc/terms/available>	"2017-06-12T06:22:57Z"^^xsd:dateTime
8	<http://rdfs.org/ns/void#sparqlEndpoint>	<http://localhost/fuseki/dspace/sparql>
9	dc:language	"en"
10	<http://purl.org/dc/terms/title>	"Discussion of the results of the observations of solar provinces made at Kodaikanal fro m 1904 to 1950"
11	dc:date	"1953-11"^^xsd:dateTime

Figure-6: Showing RDF triple results

Apache Jena Fuseki is a very powerful SPARQL endpoint query interface. It can be configured in many ways. It is capable of making queries from a large volume of RDF dataset file(s) or from a large RDF graph database. But it has a limited visualization mode. It allows visualizing the RDF data in table format or raw format. If we wish to visualize graphical mode or advanced mode of RDF data, we might go with a third party RDF browser.

8 Conclusion

This research work has presented an approach to transform the metadata of the repository into linked open data. The work attempted to explore how LOD of e-resources have been presented in a digital library environment. LOD can be regarded as an advanced information service to users who want more. LOD service is not only a traditional web service but it also provides semantic-based structured information services. Semantic web technology includes structured data and vocabulary, ontology, RDF model data, and Linked Data technology. It is a robust and complete structured data publishing system on the internet. In this research work, we have tested LOD service in DSpace digital library environment. DSpace digital library system uses PostgreSQL as a back-end database server. The PostgreSQL server contains the metadata of items of DSpace and also contains the handle number of each and every item in the repository. In the methodology section, it has been shown how all metadata of items have been transformed as RDF data in different RDF triple formats viz., RDF/XML, Turtle, N-triple, JSON-LD, etc. We have also used a SPARQL endpoint interface for RDF data querying. Some query expressions have been evaluated for accessing RDF data in different ways.

We could conclude that linked open data for digital repositories can be seen as a much wider OAI/PMH supported interface with better integration of foreign data and concepts. The DSpace digital library system can harvest metadata and full-text from other repositories as well as using OAI/PMH. The LOD support in DSpace is currently export-oriented. Only OAI/PMH can harvest all documents that change within a specified time period. To realize these into linked open data we still have to rely on vocabularies and/or conventions.

References

Ali, A.F. and Warraich, N.F.(2020).Integration of Open Educational Resources Through Linked Data: Challenges and Benefits Review. Library Philosophy and Practice (e-journal), 4530. Redtrieved from ttps://digitalcommons.unl.edu/libphilprac/4530

Apache Jena Fuseki. (n.d.). Retrieved August 12, 2020, from https://jena.apache.org /documentation/fuseki2/

Basit, I., & Hussain, H. (2019). Linked Data Applications for Content Management of Serial Publications: A Case Study of Proceedings. *Library Philosophy and Practice (e- journal), 2986*. Retrieved from https://digitalcommons.unl.edu/libphilprac/2986

Bauer, F., & Kaltenböck, M. (2012). *Linked Open Data: The essentials: A quick start guide for decision makers*. Wien: Edition mono.

Bizer, Christian, Heath, Tom and Berners-Lee, Tim, (2009). Linked Data - The Story So Far. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 5(3), 1- 22, https://EconPapers.repec.org/RePEc:igg:jswis0:v:5:y:2009:i:3:p:1-22.

Borst, T., Fingerle, B., Neubert, J., & Seiler, A. (2010). How do Libraries Find their Way onto the Semantic Web? *LIBER Quarterly*, *19*(3–4), 336. https://doi.org/10.18352/lq.7970

Das, B. and et al.(2005). XML and Its Applica tion In Information System: A Case Study. 3Rd, Convention PLANNER -2005, Assam Univ., Silchar, 10-11 Nov., 2005, Ahmedabad. Retrieved from https://ir.inflibnet.ac.in/bitstream/1944/1396/1/38.pdf

Das, R. (2014). Designing and implementing a new model of library web portal by using foss based architecture for online information service [Doctoral Dissertation]. Retrieved from http://hdl.han-dle.net/10603/133082

Das, R. (2019). Library Web Portal for Academic Libraries in India: A Case Study. Library Philosophy and Practice (e-journal), 2977.Retrieved from https://digitalcommons.unl.edu/libphilprac/2977

Das, R. & Hazra, U.(2019). An Assessment of Educational Resources through ERIC, CORE, PDFdrive and ProjectGutenberg.Library Philosophy and Practice (e-journal), 2181. Retrieved from https://digitalcommons.unl.edu/libphilprac/2181

Das, R. and et al.(2005).Designing A Digital Library with Bengali Language Support Using Unicode. 3Rd, Convention PLANNER -2005, Assam Univ., Silchar, 10-11 Nov., 2005, Ahmedabad. Retrieved from https://ir.inflibnet.ac.in/bitstream/1944/1361/1/10.pdf

Das, R. and et al.(2005). Application of DCMI In Open Source Software with SpecialReference to GSDL: A Critical Studys. 3Rd, Convention PLANNER-2005, Assam Univ., Silchar, 10-11 Nov., 2005, Ahmedabad. Retrieved from https://ir.inflibnet.ac.in/bitstream/1944/1349/1/3.pdf

Das, R.; Hazra, U.K and Mukherjee, S.(2019). e-Learning Platform in SAARC Countries. Library Philosophy andPractice (e-journal).2882. Retrieved from https://digitalcommons.unl.edu/libphilprac/2882

DSpace. (n.d.). Retrieved July 09, 2020, from https://www.lyrasis.org/DCSP/Pages/ DSpace.aspx

DSpace Metadata RDF Mapping Vocabulary (n.d.). Retrieved September 21, 2020, from http://digital-repositories.org/ontologies/DSpace-metadata-mapping/

Egusa, Y., & Takaku, M. (2020). A Case Study of Japanese Textbook Linked Open Data: Publishing a Small Bibliographic Collection from a Special Library. *International Conference on Dublin Core and Metadata Applications, 0*, 75-79. Retrieved from https://dcpapers.dublincore.org/pubs/article/view/4243/2437

Halla, M. (2013). Linked Data in Libraries: Library of Congress' Bibliographic Framework Transition Initiative. *Library Philosophy and Practice (e-journal),* 1015. Retrieved from https://digitalcommons.unl.edu/libphilprac/1015

Hazra, U., & Das, R. (2018). Evaluation of Online Learning Repositories: A Comparative Study. DESIDOC Journal of Library & Information Technology, 38(4), 259-265. https://doi.org/10.14429/djlit.38.4.12711

Hallo, M., Luján-Mora, S., Maté, A., & Trujillo, J. (2015). Current state of Linked Data in digital libraries. *Journal of Information Science, 42*(2), 117-127. doi:10.1177/0165551515594729

Heath, T., & Bizer, C. (2011). *Linked Data: Evolving the web into a global data space*. San Rafael, Calif. (1537 Fourth Street, San Rafael, CA 94901 USA): Morgan & Claypool.

Khalili, A., Loizou, A., & Harmelen, F. V. (2016). Adaptive Linked Data-Driven Web Components: Building Flexible and Reusable Semantic Web Interfaces. *The Semantic Web. Latest Advances and New Domains Lecture Notes in Computer Science,* 677-692. doi:10.1007/978-3-319-34129-3_41

Konstantinou, N., Houssos, N., & Manta, A. (2014). Exposing Bibliographic Information as Linked Open Data Using Standards-based Mappings: Methodology and Results. *Procedia - Social and Behavioral Sciences, 147*, 260–267. https://doi.org/10.1016/j.sbspro.2014.07.169

Majhi, S.C. and Das, R.(2019). Website of High Courts in India: a Webometric Study. Library Philosophy andPractice (e-journal), 2803. Retrieved from https://digitalcommons.unl.edu/libphilprac/2803

Majhi, S.C. and Das, R.(2020). Webometric Analysis of Open Access Institutional Digital Repositories in Southern Asia. Library Philosophy andPractice (e-journal), 3892. Retrieved from https://digitalcommons.unl.edu/libphilprac/3892/

Mukherjee, S. & Das, R. (2020). Integration of Domain-Specific Metadata Schema for Cultural Heritage Resources to DSpace: A Prototype Design. Journal of Library Metadata. DOI: 10.1080/19386389.2020.1834093

OpenDOAR. (n.d.). Retrieved July 09, 2020, from https://v2.sherpa.ac.uk/opendoar/

Papadakis, I., & Kyprianos, K. (2012). Providing LOD-Based Functionality in Digital Libraries. *Communications in Computer and Information Science Metadata and Semantics Research,* 13-24. doi:10.1007/978-3-642-35233-1_2

Roy, B.K. and et al.(2017). Developing Open Access Institutional Digital Repository Using Open Source Software: A Step by Step Guide. International Research: Journal of Library & Information Science,7(2). Retrieved from http://irjlis.com/wp-content/uploads/2017/08/4-IR408-621.pdf

Roy Chowdhury, A., & Das, S. (2019). Linking Library Resources to the Web through Linked Data. *Library Philosophy and Practice (e-journal), 3623*. Retrieved from https://digitalcommons.unl.edu/libphilprac/3623

Semantic Web Standards, W3C. (n.d.). Retrieved August 05, 2020, from https://www.w3.org/standards/semanticweb/

Zengenene, D. (2013). Global interoperability and Linked Data in libraries. *New Library World, 114*(1/2), 84-87. doi:10.1108/03074801311291992